

This Page Is Inserted by IFW Operations
and is not a part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

As rescanning documents *will not* correct images,
please do not report the images to the
Image Problem Mailbox.

THIS PAGE BLANK (USPTO)

02 February 2000 (02.02.2000)



Office de la propriété
intellectuelle
du Canada

Un organisme
d'Industrie Canada

Canadian
Intellectual Property
Office

An Agency of
Industry Canada

REC'D 08 FEB 2000

WIPO

PCT

*Bureau canadien
des brevets*
Certification

*Canadian Patent
Office*
Certification

La présente atteste que les documents
ci-joints, dont la liste figure ci-dessous,
sont des copies authentiques des docu-
ments déposés au Bureau des brevets.

This is to certify that the documents
attached hereto and identified below are
true copies of the documents on file in
the Patent Office.

Specification and Drawings, as originally filed, with Application for Patent Serial No:
2,259,745, on January 19, 1999, by **UNIVERSITÉ DE MONTRÉAL**, assignee of Bruno
Paquin, Ivan Brückner and Guy Tremblay, for "Generation of Oligonucleotide Libraries
Representative of Genomes or Expressed MRNAS (CDNAS) and Use Thereof".

CA00/47

4

**PRIORITY
DOCUMENT**

SUBMITTED OR TRANSMITTED IN
COMPLIANCE WITH RULE 17.1(a) OR (b)

S. D. S. D.
Agent certificateur/Certifying Officer

February 2, 2000

Date

Canada

(CIPO 68)

OPIC



CIPO

ABSTRACT OF THE DISCLOSURE

The present invention relates to oligonucleotide libraries and use thereof. In particular, the present invention relates to specific oligonucleotide
5 libraries, the oligonucleotides being of unique length, and to the use thereof. The present invention further relates to the use of these OLs in numerous biotechnological applications including a the identification and/or characterization of biological materials, clinical diagnosis (DNA/RNA level), preparative extraction of specific mRNA (and genes)
10 and genomic research/mapping.

TITLE OF THE INVENTION

GENERATION OF OLIGONUCLEOTIDE LIBRARIES
REPRESENTATIVE OF GENOMES OR EXPRESSED MRNAS (CDNAS)
AND USE THEREOF

5

FIELD OF THE INVENTION

The present invention relates to oligonucleotide librairies
and use thereof. In particular, the present invention relates to specific
oligonucleotide librairies, the oligonucleotides being of unique length , and
10 to the use thereof. The present invention further relates to the use of
these OLs in numerous biotechnological applications including a the
identification and/or characterization of biological materials, clinical
diagnosis (DNA/RNA level), preparative extraction of specific mRNA (and
genes) and genomic research/mapping.

15

BACKGROUND OF THE INVENTION

The generation of genomic DNA libraries, or cDNA
libraries and the maintenance, and handling of these libraries are critical
procedures in the field of genomics and/or biotechnology. In classical
20 libraries the relevant segments of DNA are cloned into vectors, which are
maintained and propagated in particular biological systems (in vivo).
Alternatively, libraries (in vitro) can be directly constructed from genomic
DNA or cDNA. They contain linkers at the 5' and 3' ends of the DNA
which allow PCR amplification of the library. The information stored in
25 these libraries contains repetitive sequence elements that originated from
repetitive DNA, or high copy mRNAs. This results in a significant
redundancy, which can complicate the use and the outcome of using

classical libraries. Another important feature which reduces the utility of classical libraries is their heterogeneity in size of each member of the library. Since the success of subtractive hybridization procedures is dependent upon the length, complexity, and the redundancy of the libraries, the results are particularly sensitive to the choice of method and the number of cycles performed. In fact, one must use gross hybridization conditions that can accommodate the heterogeneous length and redundancy of stored information in order to perform subtraction. Thus, the results are more "laboratory-specific" than library-specific. There thus remains a need for libraries which enable the use of hybridization conditions which are uniform, library-specific, controllable, reproducible with high resolution and are thus highly informative.

Akopyants et al. (Proc. Natl. Acad. Sci. USA 95:13108-13113) performed subtractive hybridization using bacterial DNAs digested by high-frequency restriction enzymes. The use of such a restriction enzyme tends to generate DNA fragments having a broadly similar size, about 500 base pairs. However, the uniformity is not rigorous. Furthermore, such a library of restriction fragments still contains a significant number of redundant sequences. For these reasons, the conditions of hybridization are not very well resolved. In addition, patches of short polymorphism embedded in homologous sequences are going to be missed using such a type of library.

There thus remains a need to provide a type of library which minimizes or eliminates redundant sequences, thereby enabling well resolved hybridization conditions.

The present invention seeks to meet these and other needs.

The present description refers to a number of documents, the content of which is herein incorporated by reference.

SUMMARY OF THE INVENTION

5 The present invention relates to methods for generating OLs from genomic DNAs and cDNA and to the subtraction of these libraries.

 The present invention further relates to oligonucleotide libraries which enable the use of hybridization conditions which are
10 uniform, library-specific, controllable, reproducible with high resolution and are thus highly informative.

 In addition, the invention relates to unique length oligonucleotide libraries, which minimize or eliminate redundant sequences, thereby enabling well resolved hybridization conditions.

15

BRIEF DESCRIPTION OF THE DRAWINGS

 Having thus generally described the invention, reference will now be made to the accompanying drawings, showing by way of illustration a preferred embodiment thereof, and in which:

20 Figure 1 shows A) Hybridization of random oligonucleotide mixture with template (genomic DNA/cDNA) of biological origin. High fidelity hybridization was done by using stringent hybridization conditions (as detailed in B) or mismatch free ligation procedure (as shown in C), followed by the separation of hybridized from
25 non-hybridized material and amplification of hybridized oligonucleotides. The non-hybridized material is used in the next round of rehybridization, until all complementary sequence motifs are captured. B) General

approach, in which a single oligonucleotide library is hybridized to the target DNA (1) and selected oligonucleotides are amplified by PCR (2).

C) The thermostable ligation-dependent approach, in which two oligonucleotides must hybridize side by side (1) without mismatches for the ligation to occur (the dot indicates the ligated link). As in B), the selected oligonucleotides are amplified by PCR (2).

Figure 2 shows the Subtraction between different oligonucleotide libraries (OL1 and OL2), which do contain common sequence motifs. The first step is composed of streptavidin/biotin-based partitioning of double stranded OL under denaturing conditions to separate: (i) antisense strand (which was hybridized with template) and (ii) sense strand. In the next step, antisense strand of OL1 was mixed under hybridization conditions with an excess of sense strand of OL2. The double stranded mismatch-free hybrids (OL1 antisense/OL2 sense) and single stranded OL2 sense molecules are separated from single stranded antisense-OL1 molecules. Finally, the OL1 antisense molecules are amplified and the whole procedure is repeated until none of the OL1 antisense molecules has complementary sequences to any of the OL2 sense molecules.

Figure 3 shows the validation of the Ols in accordance with the present invention. Lambda and adenovirus DNAs were dot blotted on nylon membranes and hybridized with radio-labeled oligonucleotides from the starting random pool, the OL constructed from the lambda DNA or the OL constructed from the adenovirus DNA.

Other objects, advantages and features of the present invention will become more apparent upon reading of the following non-restrictive description of preferred embodiments with reference to the

accompanying drawing which is exemplary and should not be interpreted as limiting the scope of the present invention.

DESCRIPTION OF THE PREFERRED EMBODIMENT

5 The present invention thus provides, oligonucleotide libraries having the following characteristics:

- 1) Unique length (typically 60 bp, of which the center 20 bp are representative of the library). This unique length allows clear separation of complementary *sense* from *antisense* strands
10 (biotinylated *versus* non-biotinylated strands).
- 2) Oligonucleotide libraries have a uniform number of copies for each sequence motif. Therefore, there is no differential hybridization kinetics which could originate from the presence of repetitive DNA.
- 3) As a consequence of 1) and 2), the melting profile of these OL (or
15 hybridization kinetics) has a sharp transition from double stranded to single stranded (or vice versa). This is a critical advantage in subtractive hybridization procedures.

 In one particular embodiment, the starting pool of oligonucleotides is chemically synthesized and consists of a random
20 region of a fixed length (L), flanked by a constant sequence (primer binding sites, PBS). The random oligonucleotide pool covers n copies ($n=1,2,3,\dots$) of all sequence combinations of length L, i.e. 4^L , which is a total of 10^{12} different sequence motifs for $L=20$ nucleotides. The basic length of oligonucleotides is long enough to generate unique sequence
25 motifs for a particular biological system. The complexity of the library (10^{12}) overcomes the complexity of the template (which is usually between 10^4 - 10^9). The random pool is used as starting material for

producing the biologically relevant oligonucleotide libraries (from genomic DNAs or cDNA libraries, or subtractive variants). The constant flanking sequences are used to amplify selected oligonucleotides by PCR.

The method uses a) stringent hybridization and/or b) mismatch free hybridization and ligation. High fidelity hybridization between pools of oligonucleotides and templates (genomic DNA or cDNA) is the basic mode of transfer of genomic information into OLs. A new and more efficient subtractive hybridization procedure, which differs from classical cDNA/genomic subtractions, has been developed to accommodate the features of the aforementioned OLs.

The present invention is illustrated in further detail by the following non-limiting examples.

EXAMPLE 1

15 GENERATION OF OLs, USE THEREOF IN SUBTRACTIVE
HYBRIDIZATION TO GENERATE SOLs; AND USE OF OLs OR SOLs
IN HYBRIDIZATION EXPERIMENTS

20 OLs are generated from the template DNAs. These OLs are
used in subtractive hybridization, for example between genomic or cDNA-
based library (OL1 and OL2) to make a new Subtractive Oligonucleotide
Library (SOL1/2 and/or SOL2/1)), that is/are specific for one
system/library but not for the other. Oligonucleotides isolated from such
subtractive libraries (SOL) are useful for diagnostic purposes. They can
25 a) directly serve as highly specific hybridization probes or b) they can be
tested for PCR-specific differential amplification, specific for one, but not
the other biological system.

These libraries (OL or SOL) can be hybridized to oligonucleotide chip arrays in order to obtain a specific hybridization pattern that is useful for diagnostic features: each OL produces an image which is specific for the templated DNA (genome or cDNA). A particular
5 advantage in using OL or SOL instead of genomic/cDNA libraries is that the hybridization signal is not dependent on copy number and distribution of particular sequence motifs. By comparing images of different genomes/cDNA, one can deduct which oligonucleotides are highly specific for a single genome/cDNA, and use this or these
10 oligonucleotide(s) as "genome tags". The oligonucleotides obtained can also be used for specific diagnostic PCR.

Ols or SOLs can be inferred from two biologically relevant systems, like mammalian cells, to detect fine differences in cell cycle, tissue status, viral infection, age/development status etc.

15 **Experimental facts:**

To validate our approach, we first tested the procedure with a simple system. We generated OLs using lambda phage and adenovirus genomes as templates, separately. The complexity of these genomes is around 10^4 , which corresponds to a small fraction of the total complexity
20 of the starting random pool ($1/10^8$). Lambda and adenovirus genomes do not have any 20-mer motifs in common, which means that an OL made from lambda DNA should not cross-hybridize to the adenovirus DNA, and *vice versa*. The starting mixture of random oligonucleotides is such that it should contain about 10 copies of each possible 10^{12} different
25 sequences. As shown in Fig. 3A), when the random pool of the oligonucleotides is radio-labeled, both genomes are hybridizing to the probes. However, after construction of high fidelity OLs, the hybridization

signals become specific to the genome from which the OL was inferred (Fig. 3). Therefore, OLs are indeed reflecting the original template.

Second, artificial mixture of template DNAs (lambda and adenovirus) is made in different ratios to mimic conditions of subtractive hybridization. Having lambda OL on the one hand and these different mixtures on the other hand, we are proving that we could subtract the lambda-specific oligonucleotides, and isolate the adenovirus-specific ones. We are then planning to test these concepts on more complex systems, like bacterial genomes (complexity of about 10^6) and, ultimately, mammalian cells (complexity of about 10^9).

Although the present invention has been described hereinabove by way of preferred embodiments thereof, it can be modified, without departing from the spirit and nature of the subject invention as defined in the appended claims.

WHAT IS CLAIMED IS:

1. A method of generating an oligonucleotide-based library which originates from a chosen biological material, comprising:

- 5 a) generating a random oligonucleotide mixture, wherein said oligonucleotides are of a unique length;
- b) hybridizing said random oligonucleotide mixture of a) with a nucleic acid template of biological origin under hybridization conditions which enable the formation of duplexes, while minimizing or abrogating
- 10 mismatches;
- c) separating said duplexes from non-duplexed material;
- d) amplifying said hybridized oligonucleotides;
- e) rehybridizing non-hybridized material with said template;
- f) repeting said steps a) to e) until all oligonucleotides that
- 15 can hybridize have been captured.

2. The method of claim 1, wherein said template is genomic DNA or cDNA.

- 20 3. The method of claim 1 or 2, wherein said hybridization conditions are carried out using one of: stringent hybridization conditions, mismatch-free ligation of random oligonucleotides, type 5' primer(N)₇-₁₀ 3' and 5'(N)₇₋₁₀ primer 3', using thermostable ligases, or a combination thereof, mismatch-free ligation of random oligonucleotides, type 5'
- 25 primer(N)₇₋₁₀ 3' and 5'(N)₇₋₁₀ primer 3', using thermostable ligases, or combination thereof.

4. The method of claim 1, 2, or 3, wherein said biological material is selected from genomic DNA, and cDNA.

5. A method of subtracting between different oligonucleotide libraries which contain common sequence motifs comprising:

a) a partitioning of double stranded OL under denaturing conditions to separate: (i) antisense strands hybridized with template and (ii) sense strands;

b) annealing said antisense strand of OL1 with an excess of OL2 sense strand, under hybridization conditions;

c) partitioning of double stranded mismatch-free hybrids (OL1 antisense/ OL2 sense) and single stranded OL2 sense molecules from single stranded antisense OL1 molecules; and

d) amplification of OL1 antisense molecules and repetition of steps a) to d) until OL1 antisense does not have complementary sequences to OL2 sense.

6. The method of claim 5, wherein said partitioning is carried out using streptavidin/biotin.

Fig 1A

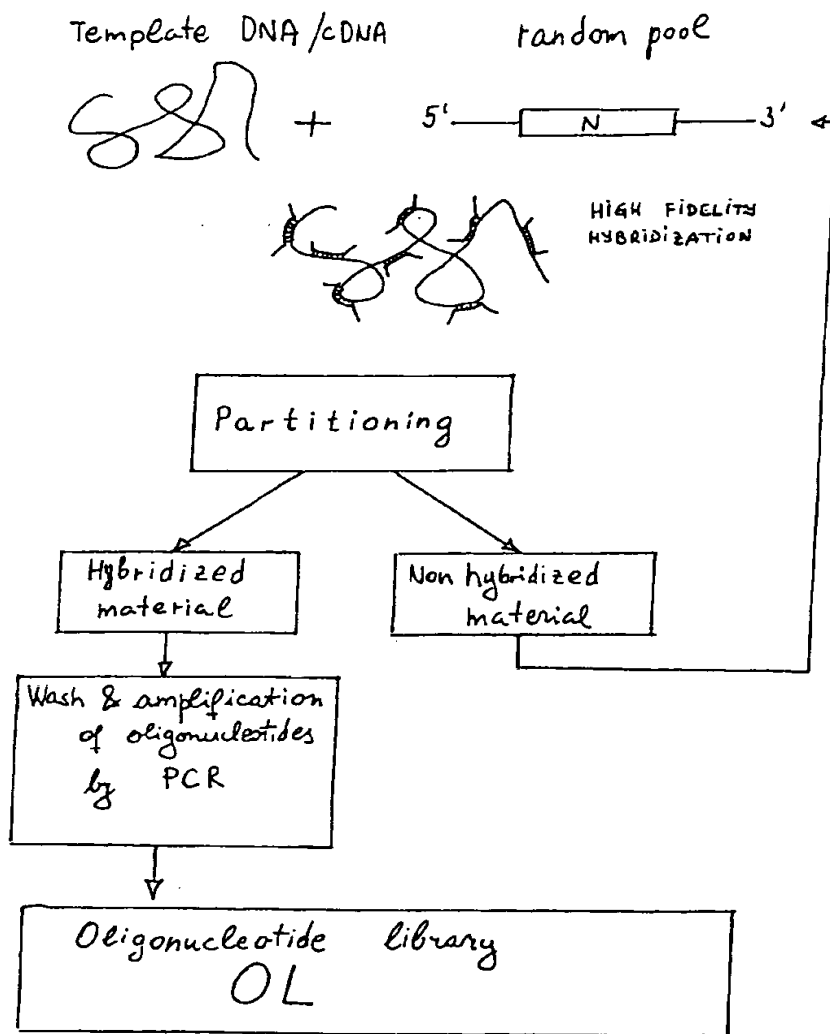


Fig. 2

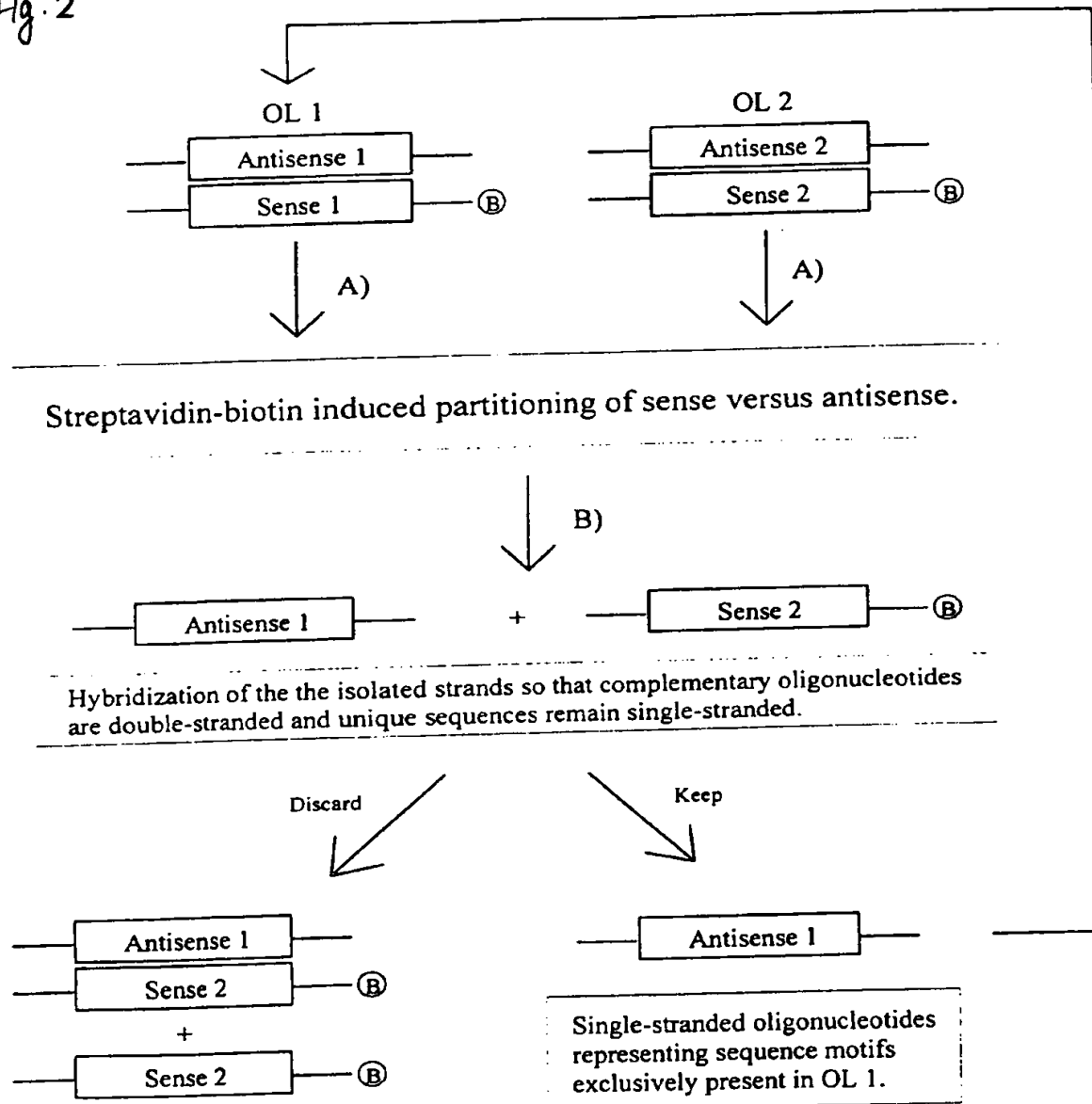
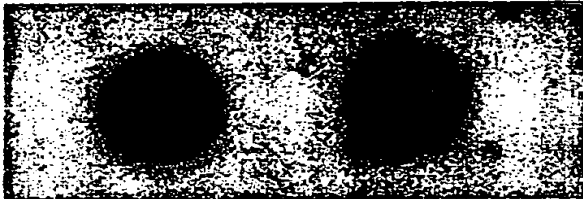
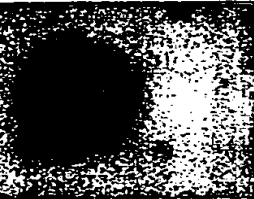



Fig. 3

Probes	Target DNA	
	Lambda	Adenovirus
Random oligonucleotides		
Lambda OL		
Adenovirus OL	